data1

来自NIC[1] dataset1

标签(200)：1(70个)、2(60个)、3(70个)

视图：X1(0-4.2089)、X2(二元变量0、1)

维度:X1(1000\*200)、X2(5000\*200)

data2

来自NIC[1] dataset2

标签(200)：1(70个)、2(60个)、3(70个)

视图：X1(0-7.0061)、X2(二元变量0、1)

维度：X1(1000\*200)、X2(5000\*200)

data3

来自NIC[1] mESC

标签(77)：1(13个)、2(64个)

视图：DNA(二元变量0、1)、RNA(0-13.8720)

维度：DNA(15000\*77)、RNA(5500\*77)

data4

来自JSNMF[2] brain5k\_10k

视图与维度：

X1(5000\*3057)、

X2(10000\*3057)

标签（3057）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 标签 | 个数 | 标签 | 个数 | 标签 | 个数 |
| 1 | 108 | 7 | 132 | 13 | 111 |
| 2 | 324 | 8 | 93 | 14 | 65 |
| 3 | 234 | 9 | 77 | 15 | 51 |
| 4 | 222 | 10 | 65 | 16 | 184 |
| 5 | 202 | 11 | 563 | 17 | 27 |
| 6 | 190 | 12 | 176 | 18 | 85 |
|  |  |  |  | 19 | 148 |

data5

来自JSNMF[2] h3k4me1\_5k\_10k

视图与维度：

X1(0-6.4398) (5000\*8637)

X2(0-5.7862) (9999\*8637)

标签（8637）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 标签 | 个数 | 标签 | 个数 | 标签 | 个数 |
| 1 | 1663 | 9 | 0 | 17 | 409 |
| 2 | 501 | 10 | 627 | 18 | 127 |
| 3 | 915 | 11 | 0 | 19 | 574 |
| 4 | 421 | 12 | 470 | 20 | 389 |
| 5 | 239 | 13 | 78 | 21 | 512 |
| 6 | 280 | 14 | 98 | 22 | 107 |
| 7 | 155 | 15 | 65 |  |  |
| 8 | 132 | 16 | 875 |  |  |

data6

来自JSNMF[2] h3k4me3\_5k\_10k

视图与维度：

X1(0-6.2371) (5000\*2684)

X2(0-6.2222) (9903\*2684)

标签（2684）



data7

来自JSNMF[2] h3k9me3\_5k\_10k

视图与维度：

X1(0-6.2600) (5000\*8176)

X2(0-5.6423) (10000\*8176)

标签(8176)



data8

来自JSNMF[2] h3k27ac\_5k\_10k

视图与维度：

X1(0-6.4373) (5000\*7262)

X2(0-5.9732) (10000\*7262)

标签(7262)



data9

来自JSNMF[2] h3k27me3\_5k\_10k

视图与维度：

X1(0-6.2174) (5000\*4307)

X2(0-6.4327) (10000\*4307)

标签(4307)



data10

来自JSNMF[2] kidney\_5k\_10k

视图与维度：

X1(0-7.8318) (5000\*4753)

X2(0-8.3522) (10000\*4753)

标签(4753)



data11

来自ATAC\_RNA\_FIT[3]

5个mat

distal\_index(491437)(0、1)

single\_cell: 数据(2034\*491437)(0-42)

type(name)(2034): HSC(347)、MPP(142)、LMPP(160)、CMP(502)、CLP(78)、MEP(138)、GMP(216)、pDC(141)、Mono(64)、UNK(60)、GMP3high(74)、GMP2mid(44)、GMP1low(68)

scRNAseq\_data: 数据(12558\*14432)(0-838)

cell Labels(14432): hsc(2268)、cmp0(2508)、cmp1(626)、cmp2(1320)、gmp(1096)、mono(129)、cd34(6485)

gene Labels(12558)

RNA\_bulk(32738\*19)(0-20.5236) col(19)细胞类型 row(32738)基因

col: UNK、GMP1、GMP2、GMP3、MKP、pDC、Bcell、CD4、CD8、NK、CLP、Ery、LMPP、CMP、GMP、HSC、MEP、Mono、MPP

cqn(17\*491437)(0-29.8599)

cqnlabel(17):HSC、MPP、LMPP、CMP、GMP、MEP、Mono、CD4、CD8、NK、NKT、Bcell、CLP、Ery、UNK、pDC、Mega

data12

来自GEO[4](GSM3271044 and GSM3271045)

RNA组学 数据txt文件(49584 13893 7036187) cell\_label(13893) gene(49584)

ATAC组学 数据txt文件(252741 13395 9448526) cell\_label(13395) peak(252741)

data13

来自scAI[5] sciCAR

ATAC(141290\*3260) cell\_label(3260) features(141290)

RNA(47168\*3260) cell\_label(3260) features(47168)

data14

来自scAI[5]

Label1: 1(100)、2(100)、3(100)、4(100)、5(100)

Label2: 1(100)、2(100)、3(100)、4(200)

X1(2000\*500) (0-4.6295)

X2(5000\*500) (0、1)

data15

来自scAI[5]

Label1: 1(100)、2(100)、3(100)、4(200)

Label2: 1(100)、2(100)、3(100)、4(100)、5(100)

X1(2000\*500) (0-4.5609)

X2(5000\*500) (0、1)

data16

来自scAI[5]

Label: 1(50)、2(50)、3(50)

X1(2500\*150) (0\28-2404348)

X2(5000\*150) (0、1)

data17

来自scAI[5]

Label: 1(50)、2(50)、3(50)、4(50)、5(50)

X1(2500\*250) (0-6.3908)

X2(5000\*250) (0、1)

data18

来自scAI[5]

Label: 1(100)、2(1000)

X1(1000\*1100) (0-4.5565)

X2(5000\*1100) (0、1)

Reference

1. Wu W, Zhang W, Ma X. Network-based integrative analysis of single-cell transcriptomic and epigenomic data for cell types. Brief Bioinform. 2022 Mar 10;23(2):bbab546.
2. Ma Y, Sun Z, Zeng P, Zhang W, Lin Z. JSNMF enables effective and accurate integrative analysis of single-cell multiomics data. Brief Bioinform. 2022 Apr 4:bbac105.
3. Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY, Greenleaf WJ. Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of Human Hematopoietic Differentiation. Cell. 2018 May 31;173(6):1535-1548.e16. <https://ars.els-cdn.com/content/image/1-s2.0-S009286741830446X-mmc4.zip>
4. https://www.ncbi.nlm.nih.gov/geo/
5. Jin S, Zhang L, Nie Q. scAI: an unsupervised approach for the integrative analysis of parallel single-cell transcriptomic and epigenomic profiles. Genome Biol. 2020 Feb 3;21(1):25.